



Intentional joint agency: shared intention lite

Elisabeth Pacherie

► To cite this version:

Elisabeth Pacherie. Intentional joint agency: shared intention lite. *Synthese*, 2013, 190 (10), pp.1817-1839. 10.1007/s11229-013-0263-7 . ijn_00872152

HAL Id: ijn_00872152

https://hal.science/ijn_00872152

Submitted on 11 Oct 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intentional joint agency: shared intention lite

Elisabeth Pacherie
Institut Jean Nicod
IEC-ENS, EHESS, CNRS, Paris

Abstract

Philosophers have proposed accounts of shared intentions that aim at capturing what makes a joint action intentionally joint. On these accounts, having a shared intention typically presupposes cognitively and conceptually demanding theory of mind skills. Yet, young children engage in what appears to be intentional, cooperative joint action long before they master these skills. In this paper, I attempt to characterize a modest or 'lite' notion of shared intention, inspired by Michael Bacharach's approach to team-agency theory in terms of framing, group identification and team reasoning. I argue that the account of shared intentions this approach yields is less cognitively and conceptually demanding than other accounts and is thus applicable to the intentional joint actions performed by young children. I also argue that it has limitations of its own and that considering what these limitations are may help us understand why we sometimes need to take other routes to shared intentions.

Keywords: joint action; joint goal; shared intention; team reasoning; group identification; frames; Michael Bratman; Michael Bacharach.

Intentional joint agency: shared intention lite

1. Introduction

Joint action takes a multitude of forms and occurs in a variety of systems and at multiple levels of complexity. What seems to distinguish human joint action from the collective behavior observed in many other social species is its intentional character: human joint action is very often an intentional cooperative activity where agents' actions are consciously subordinated to intentionally shared goals. Philosophers have developed theories of shared intentions to account for what makes a joint action intentionally joint. While there is much disagreement among them concerning how best to construe shared intentions, their accounts tend to focus on joint actions performed by human adults, on the processes involved in making decisions about whether to act together, and on the processes involved in high-level planning. As a result, these philosophical accounts suffer from two important limitations. First, they are ill suited to joint actions performed by young children who have yet to develop the psychological structures they deem necessary for joint actions. Second, as their focus is on the stages that precede actual engagement in joint action, they have little to say on what makes it possible for agents to successfully carry out a joint action and thus realize their shared intention.

In contrast, cognitive psychology studies of joint action typically focus on the perceptual, cognitive, and motor processes that enable individuals to coordinate their actions with others online and developmental psychology studies attempt to identify and characterize the stages through which infants and young children progressively become more adept and active partners in joint activities. There is thus a gap between the dominant philosophical approaches and the empirical investigations of joint agency. As explained by Butterfill and Sebanz (2011), two broad strategies for dealing with this situation can be discerned. One involves breaking away from philosophical accounts of joint actions in terms of shared intentions and arguing that the joint actions in which young children engage do not involve shared intentions but something weaker – for instance, shared goals – and compatible with the more primitive levels of social understanding and communication in infancy (Apperly and Butterfill 2009; Butterfill

2012; Michael, 2011; Seemann, 2011). The second strategy involves retaining the link between shared intention and joint intentional action, while offering a potentially less demanding account of what counts as shared intention (Blomberg, 2011; Pacherie, 2011; Tollefsen 2005). As Butterfill and Sebanz (2011) also point out, these two strategies should be seen as complementary to each other rather than exclusive and the second strategy should not be seen as aiming to eliminate sophisticated notions of shared intention altogether.

In this paper, I will pursue the second strategy and attempt to characterize a modest or 'lite' notion of shared intention, less cognitively demanding than what the analyses proposed by leading philosophical accounts suggest and constituting a plausible basis from which more sophisticated forms of shared intentions can, under certain conditions, emerge. I start with a brief discussion of what is required for joint actions to be intentionally joint. I then introduce Michael Bratman's account of shared intentions. While happy to admit that shared intentions can sometimes take the form proposed by Bratman, I will give some reasons for doubting that his account can apply to all intentional joint actions. In particular, I am skeptical that all intentional joint actions require the sophistication in ascribing propositional attitudes that Bratman's account appears to demand. To motivate this skepticism, I'll turn to developmental psychology work on joint action in young children and discuss empirical evidence that young children engage in what appears to be intentional joint action despite lacking this conceptual sophistication.

The 'lite' account of shared intention I will present is inspired in part by certain recent developments of game theory, known as theories of team agency, that challenge the individualistic rationality assumptions of standard game theory (Bacharach 2006; Sugden 1993, 2003; Gold and Sugden 2007, 2008). My primary interest here will be with Bacharach's view that team formation and therefore team agency is the result of framing. I will try to show how a 'lite' account of shared intentions can be derived from these ideas. I will also argue that the cognitive toll one has to pay to take this route to shared intention is more modest than the cognitive toll demanded to take Bratman's route. I do not want to claim, however, that the team-agency approach is the only route to shared intention. Rather, I think that considering some of the limitations of this approach can help us understand why we sometimes need to take other routes.

2. *Intentional joint agency: what is at stake?*

Through actions we bring about changes in the environment. Actions are typically described in terms of their outcomes or effects. As a single action can have a wide range of outcomes, it is in principle describable in a host of ways. As a first pass, one may want to say that an event qualifies as a joint action only if it is describable as the common outcome of what several agents did; that is, none of the agents involved did on their own bring about this outcome. Let us call this the *common outcome requirement*.

Not all the changes we bring about in our environment count as actions. When we sneeze and thus release millions of microbes in the air, we certainly have an effect on our environment. Although there is a sense in which sneezing is something we do, few consider sneezing a genuine action. Davidson (1980, essay 3) famously argued that for an event to qualify as an action it must be something the agent does that is intentional under some description. Many philosophers have agreed with him that there was an important tie between action and intention, although determining the exact nature of this tie has proven a difficult challenge and given rise to lively debates.

The challenge is not confined to individual action; indeed, many would think that in joint action the problem strikes with a vengeance. Minimally, we want to exclude cases where a common outcome is caused by doings we do not wish to count as actions. To do so, we should also require that the agent's individual doings that together bring about the common effect be actions, that is, in Davidson's parlance, that they be intentional under some description. Let us call this second requirement the *individual intentional action requirement*.

Clearly though, actions that satisfy these two requirements do not yet constitute joint actions. Suppose that Pierre's action is intentional under the description washing the dishes, and Marie's action is intentional under the description taking a shower and that as a common effect of both their actions the hot water tank is emptied. Here both the common outcome requirement and the individual intentional action requirement are satisfied, and yet emptying the hot water tank isn't intuitively a joint action. In particular, satisfying these two requirements is compatible with neither Pierre nor Marie intending to empty the hot water tank. To exclude such cases, we should also

require that the agent's actions aim at bringing about the same effect. Let us call this third requirement the *common goal requirement*. What the common goal requirement imposes is that the common outcome requirement and the individual intentional action requirement be related in a particular way: in other words, there must be a certain outcome *O* to which several agents actually contribute and the actions of each of the agents involved must be intentional under the description bringing about outcome *O*. It is important to note that the sense in which this requirement imposes that the agents have a common goal is simply distributive: the fact that the goal is common to the agents is not part of the description under which each intends their own action.

Actions that satisfy these three requirements still can't be considered joint actions, rather they correspond to what Butterfill (2012) calls plural activities. For suppose now that Pierre and Marie each intends to empty the hot water tank and that their actions indeed together contribute to emptying the hot water tank. This is compatible with neither Marie nor Pierre realizing that the other is pursuing the same goal, is at home, or even exists – for instance, we could imagine that Pierre is just back from a trip and that Marie is a new housemate who arrived while he was absent.

What Butterfill suggests is still missing is some form of coordination among the agents' actions. Let us call this fourth requirement the *action coordination requirement*. But, as Butterfill also explains, coordination can take very different forms. It can be achieved hormonally (as in the case of ants) or through psychological mechanisms. While some of the psychological mechanisms that ensure action coordination are under intentional control, others aren't. Various forms of emergent coordination, where coordinated behavior occurs due to perception–action couplings that make multiple individuals act in similar ways, would fall under the latter category.¹

While the addition of the action coordination requirement allows us to capture a notion of joint action, joint action so characterized still falls short of what most philosophers are interested in. To put it in a nutshell, they are interested in joint actions that are intentional as joint. The action coordination requirement doesn't give us that insofar as it allows joint actions to involve only involuntary forms of emergent coordination.

¹ For a discussion of various forms of emergent coordination, see Knoblich & Sebanz (2008) and Knoblich, Butterfill & Sebanz (2009).

One will come closer to what philosophers are interested in if one adopts a modified version of the action coordination requirement and allows joint actions to count as intentionally joint only if voluntary coordination processes are at work. To cash this out we can appeal to Butterfill's notion of a shared goal, where shared goals are characterized in terms of their function, which is to "coordinate multiple agents' goal-directed activities around an outcome to be achieved as a common effect of their efforts" (2012: 37). Importantly, in contrast to other forms of coordination, coordination by shared goals is voluntary and dependent on which goals agents' actions are directed to. We can call this modified version of the action coordination requirement the *intentional action coordination requirement*.²

Many philosophers appeal to shared intentions rather than shared goals in order to capture what makes joint actions jointly intentional. As we will see in the next section, accounts that appeal to shared intentions tend to be more cognitively costly than the kind of account Butterfill proposes. This difference in costs relates to two differences between shared goals as conceived by Butterfill and shared intentions. First, as Butterfill explains, while shared intentions and shared goals both exist to coordinate activities, shared intentions, at least on Bratman's understanding, have further functions that shared goals lack. Shared intentions also serve to structure bargaining over what to do and about how to divide tasks among agents and to coordinate the agent's activities over the long term. Second, while sharing an intention requires awareness that the intention is shared, no such requirement applies to shared goals. It is possible, according to Butterfill, to have a shared goal without knowing or even suspecting that one does: "agents can have, and act on a shared goal without understanding their actions as comprising anything more than a plural activity" (2012: 42).

On this characterization of shared goals and of intentional joint action as involving shared goals, the individual actions of two agents treating each other as social tools could together constitute an intentional joint action. Suppose that Ahmed's situation is as follows: his goal is to bring about outcome *O*; he realizes that Malika has the same goal; he also realizes that he could use her as a social tool to achieve his goal more easily

² It is important to note that in using Butterfill's notion of a shared goal in order to characterize a weaker form of intentional joint action, I go beyond what Butterfill (2012) commits himself to. His objective in this paper is to provide an account of a kind of joint action that does not involve shared intentions. He takes no stand on whether or not the kind of joint action he is interested in qualifies as intentional.

and that to use her efficiently as a social tool he must coordinate his actions with hers; as a result, he indeed coordinates his actions with hers. Suppose that Malika's situation is completely symmetric to Ahmed's. Suppose neither of them realizes (or cares) that the other is simply using them as a social tool. Ahmed and Malika together bringing about outcome *O* would meet the *intentional action coordination requirement* and thus qualify as an intentional joint action in this weak sense.

In contrast, philosophers who appeal to shared intentions are trying to capture a stronger notion of intentional joint action. For a joint activity to be a joint intentional action in this strong sense, the individuals who engage in this activity must think of its goal not just as bringing about outcome *O*, but as bringing about outcome *O together*. To help clarify the difference between a weaker and a stronger notion of joint intentional action, let me introduce a distinction between agent unmarked specifications of goals – e.g., *to move the table* – and agent marked specifications of goals – e.g., *I move the table* or *we move the table*. Suppose I want to show my strength, then the goal of my action will not just be to move the heavy table but to move it on my own. My goal is first-person singular agent marked. Suppose all I care about is that the table be out the way, my goal is that the table be moved. If others happen to share this goal, I may coordinate my movements with theirs and succeed more easily with their help. Then, our moving the table together will be the joint means of satisfying the agent unmarked goal we happen to share. This action constitutes a joint intentional action in the weaker sense. In contrast, for there to be a joint intentional action in the stronger sense, the agents should have a first-person plural agent marked goal (a we-goal or joint goal, for short) – *we move the table* -- and their sharing this we-goal should serve to coordinate their goal directed activities. In other words, agents engaged in a strongly intentional action do not just intend to achieve a common outcome, they intend to achieve it jointly. Part of what it means is that they should view their own actions as contributions to a we-goal rather than viewing theirs and the other agents' actions as contributions to some agent unmarked goal they happen to share (whether they are aware of it or not). We can call this final requirement the *joint goal requirement*.

It is difficult to further articulate what this last requirement involves at this point without already committing oneself to a particular view of shared intentions. Let me mention, however, two opposite pitfalls one should strive to avoid. The first danger lies

in too weak a characterization. Suppose, Pierre and François are both convinced that the other is a thief. Each wants to take the other to the police station. While on their way to the police station, each is ready to catch the other by the collar and drag him to the police station in case he tries to run off. In a sense, one could say that each has a goal expressible as "we go to the police station together", yet, intuitively, their action is not a jointly intentional action in the strong sense. In particular, Pierre and François presumably do not think of their joint activity as a cooperative activity, whose aim is not just that a certain outcome be achieved but that it be achieved together. The goal of each could as adequately be rephrased as "I take you to the police station [whether you are willing to go or not]". Rather, for an action to be jointly intentional in the strong sense, the agents should think of it as a collaborative activity. But then a second danger looms. For if we say that for us to share an intention to A is for us to intend to collaborate to achieve A, then this appeal to shared intentions does not appear to have explanatory value. Rather, we seem to be running in a circle. If we are to appeal to shared intentions to explain what is special about strongly intentional joint actions, the challenge is therefore to provide an informative analysis of shared intentions. This does not mean necessarily that such an analysis has to be reductive; however, to be of value a non-reductive analysis should strive to identify the irreducible core of the notion and characterize it in an informative way.

To recap, we have considered six requirements: (1) the common outcome requirement; (2) the individual intentional action requirement; (3) the common goal requirement; (4) the action coordination requirement; (5) the intentional action coordination requirement; and (6) the joint goal requirement. It is necessary but not sufficient that requirements (1)-(3) be satisfied for there to be a joint action. Actions that also meet requirement (4) qualify as joint actions, but not necessarily intentional joint actions, as requirement (4) allows that action coordination among agents be achieved by emergent coordination processes alone. With the addition of requirement (5), we can characterize a weak form of joint intentional action. Actions that meet requirements (1)-(5) are intentional *qua* joint only in the sense that agents intentionally coordinate the means they use to achieve a common goal. In contrast, joint actions that satisfy requirement (6)

are joint in a stronger sense, for there the joint-ness of their actions is not just a means to a goal, it is part of the goal itself.³

Butterfill's project is to give a simple account of joint action that can help us characterize the joint actions performed by young children. His interest is in joint actions that meet requirements (1)-(5). While I find his project very congenial, I think there is room for an account that is slightly more ambitious than Butterfill's in that it captures the joint goal requirement and yet is less cognitively costly than traditional philosophical accounts of shared intentions. In other words, I want to argue that it is possible to have one's cake and eat it!

But first, let us examine in what sense philosophical accounts of shared intentions may be overly cognitively costly.

3. Heavyweight shared intentions

Philosophical accounts of shared intentions⁴ are attempts to cash out what it takes for agents to act in a jointly intentional manner. These accounts all agree that joint actions are more than mere summations of individual actions and that the intentionality in joint action cannot reduce to the intentionality of the individual actions that together contribute to the joint action. They all agree therefore that something more is needed. But what more is needed? This is where disagreements start. Some philosophers hold that shared intentions involve a *sui generis* attitude, such as a *we*-intention (Searle 1990) or a participatory intention (Kutz 2000); others identify the mark of shared intentionality at the level of the subject, holding that shared intentions can only be attributed to a plural subject (Helm, 2008) or emphasize the patterns of interlocking obligations or commitments imposed on participants in joint action (Gilbert 1992, 2009); and others, again, hold that shared intentions involve a specific pattern of

³ Note that for there to be intentional joint actions in the strong sense implied by requirement (6), it isn't always necessary that requirements (4) or (5) be satisfied. For instance, Searle (1990) considers as an instance of intentional joint action the case where business school graduates all get together on graduation day and form a pact to the effect that they will all go out together and help humanity by way of each pursuing his own selfish interests and not cooperating with the others. This case satisfies requirement (6), but neither requirement (4) nor requirement (5). It is also possible that a joint action meet requirements (4) and (6) but not (5), as when all the action coordination needed to achieve the joint goal is taken care of by emergent coordination processes.

⁴ Different authors use different terminologies, speaking of shared intentions, collective intentions, joint intentions or *we*-intentions. Here I use these labels interchangeably, unless otherwise stated.

interdependence among individual intentions whose contents interlock in specific ways (Bratman, 1992; 2009). As Butterfill (2011) points out, despite their differences, most leading accounts of shared intentions make similar demands on agents who act on a shared intention: (1) Agents must be aware that they are not acting individually and think of their action as their contribution to a joint goal; (2) agents must be aware of (at least some of) the other contributing agents as intentional agents; (3) agents must act in part because of their awareness of other's agency and of the joint-ness of their goal; and (4) agents must be aware of at least some of the other agents' attitudes concerning the joint action. The cognitive cost of philosophical accounts of shared intentions is linked to these demands.

In the remainder of this section I will focus on Bratman's account, first because it is, in my opinion, one of the most convincing accounts on the market and, second, because its explicitness makes it possible to see very clearly what kind of abilities it requires of agents to share intentions.

The aim of Bratman's account is to provide a set of sufficient conditions for shared intentions using the conceptual and normative resources of his planning theory of individual agency. His approach to shared intentions is thus a constructivist approach: he thinks we can account for the joint-ness of shared intentions if we construe them as structures of suitably inter-related intentions and other attitudes of participants, where the component intentions of the individuals belong, *pace* Searle, to the ordinary brand of intentions, only with special and distinctive contents and interrelations. He also thinks, *pace* Gilbert, that the social norms that apply to and guide shared intentions can be shown to emerge from the norms that govern individual planning agency.

Bratman (2009) proposes that shared intention involves the following main building blocks:

- (1) Intentions on the part of each in favor of the joint activity.
- (2) Interlocking intentions: each intends that the joint activity go in part by way of the relevant intentions of each of the participants.

- (3) Intentions in favor of meshing subplans: each intends that the joint activity proceeds by way of subplans of the participants that are co-realizable and can be consistently agglomerated.
- (4) Disposition to help if needed: given that the contribution of the other participants to the joint activity is part of what each intend, and given the demands of means-end coherence and of consistency that apply to intentions, each is under rational pressure to help others fulfill their role if needed.
- (5) Interdependence in the persistence of each participant's relevant intention: each continues to intend the joint activity if and only if (they believe) the other participants continue to so intend.
- (6) Joint-action-tracking mutual responsiveness: each is responsive to each in relevant subsidiary intentions and in relevant actions in a way that tracks the joint action.
- (7) Common knowledge among all participants of all these conditions.

I won't comment extensively on each of these conditions here, but let me offer some remarks. First, although the first condition refers to the joint action, it is not the condition that is supposed to account for the joint-ness of the activity in the strong sense in which we are interested. If it were, the account would be circular. Indeed, Bratman takes care to point out that the concept of a joint activity that figures in the contents of the intentions in (1) should be understood in a way that is neutral with respect to shared intentionality. Instead, it is condition (2) that is the most central to Bratman's account of shared intentionality. It is the fact that for each participant, the content of their intention refers to the role of the intentions of other participants that, for Bratman, captures the intentional joint-ness of their actions. Conditions (3), (4) and (6) can be derived from condition (2) taken together with the norms already associated with individual planning and acting. It should also be remarked that conditions (1) and (2) both violate the own action condition, i.e. the constraint that one can only intend one's own actions, with condition (1) ranging over others' actions and condition (2) over others' intentions. However, Bratman (1992) argues that conditions on intending *that* are weaker than conditions on intending *to*. Roughly put, for one to intend that *A*, it is not necessary that one suppose this intention will lead one to do *A*, it suffices that one

suppose it will lead one to do something that has an influence on whether or not *A* obtains. Condition (5) clarifies the nature of this influence: each believes that his or her intention controls the intentions and actions of others by way of its support of the persistence of the other's relevant intentions.

Bratman's ambition in giving this account is avowedly modest. His interest is in cases of small-scale shared intentional agency in the absence of asymmetric authority relations. He purports to offer sufficient conditions for shared intentions rather than necessary and sufficient conditions, allowing therefore for possible alternative constructions and leaving open the possibility that shared intention is multiply realizable.

As an account targeting small-scale shared intentional activity Bratman's account has much to recommend it. By conceiving of shared intentions as an interlocking web of intentions of individuals, it moves away from the classical reductive analyses of collective action, since it maintains that the crucial link among the attitudes of agents involved in joint activity is not just a matter of mutual belief or mutual knowledge. At the same time, the account is metaphysically and normatively parsimonious. On Bratman's view, it is not necessary to introduce a new *sui generis* kind of mental state; ordinary intentions do the trick. It is not necessary to introduce new supra-individual entities to whom shared intentions would be ascribed; instead, it is sufficient that we ascribe a set of intentions, *seriatim*, to individual human agents. It is not necessary either to see the social normativity characteristic of shared intentions as a basic, non-reducible form of normativity; rather this social normativity emerges from the normativity already associated with individual planning agency.

Yet, there are reasons to think that, even leaving aside cases where asymmetric authority relations are present, alternative construals of shared intentions will be needed. One important limitation to the scope of Bratman's constructivist account comes from the fact that while the materials come cheap (it makes do with ordinary intentions and with the normativity already present in individual planning agency), their assemblage is costly and demands cognitively sophisticated agents. As pointed out by Butterfill (2012), Pacherie and Dokic (2006) and Tollefsen (2005), Bratman's analysis requires that the participating agents have concepts of mental states, since each participant should represent that the other participants have intentions and other

attitudes relevant to the joint activity. It requires not just first-order but also higher-order theory of mind abilities insofar as sharing intentions requires in addition that each agent have intentions about their own and others' intentions participant (see Bratman's condition 2 above) and that this be common knowledge among them (Bratman's condition 7). Bratman's account would therefore be inapplicable to joint actions performed by creatures lacking or having yet to develop the sophisticated cognitive skills it requires. As we will see in the next section, there is empirical evidence that young children do engage in what appears to be intentional joint action before they master these skills. This in itself does not invalidate Bratman's account since it purports to give sufficient but not necessary conditions for shared intentions and thus leaves open the possibility that the functional role of shared intentions be realized in different ways. Yet, if indeed young children engage in joint action, this puts pressure on us to propose an alternative account of shared intentions. The task here may not just be to give an alternative set of conditions that could be met by creatures simpler than psychologically sophisticated adult humans, but also to consider an alternative functional characterization of shared intentions.

A preliminary question then is whether we should go along with Bratman's functional characterization of shared intentions. Bratman's main interest is in what he calls planning agency as opposed to mere purposive agency and his functional characterization of shared intentions is strongly influenced by his view of human agents as planning agents. We are planning agents in that we frequently settle in advance on more or less complex plans and in so doing commit ourselves to future conduct in ways that support complex forms of organization and coordination both over time and with others. As a result, Bratman's conception of intentions is as elements of larger plans and his focus is on the function of intentions in insuring the consistency and stability over time of these larger plans. In contrast, Bratman is much less concerned with the roles intentions may play in the selection, implementation, monitoring and control of the actions that need to be carried out for the intention to be realized. In other words, he's much less interested in the role of intentions in action planning as understood in research on motor cognition.⁵ These are presumably questions he would take to be the province of a theory of purposive agency. However, Bratman's category of purposive

⁵ See, for instance, Jeannerod (1997).

agency is highly generic and encompasses all agents, human and non-human, "who pursue goals in light of their representations of the world" (Bratman 1999: 5). While the spider's activity when spinning its web and my activity when preparing a cup of coffee are both purposive activities in this generic sense, there is certainly room for distinctions among their respective forms of purposiveness. One intuitive way of characterizing what distinguishes my purposive activity from the spider's is to say that my activity, but not the spider's, is guided by my conscious intention to prepare a cup of coffee. In other words, my purposive agency, but not the spider's, would qualify as intentional agency. Here it seems that for talk of intention to be licensed, it is not necessary that my intention to prepare coffee be part of some larger plan.

Theories of intention may be developed as part of different projects. Bratman's project of characterizing planning agency is one such project; but a different project can be to characterize a specific form of purposive activity and to capture what is special about the goal-directed activities of agents whose goals are conscious and whose conscious representations of their goals play a role in action guidance. Indeed, it is very likely that being capable of this specific form of purposive agency is a prerequisite for planning agency. Hence, it seems that the two projects are not exclusive but complementary and that a 'lite' theory of intentions focusing on their role in the control of certain forms of goal-directed behavior may be a necessary step towards a more heavyweight theory of intentions focusing, as Bratman's does, on their role in insuring the stability and consistency of long-term plans.

Similarly, instead of simply having a generic category of purposive shared activities where the shared activities of a swarm of bees are lumped together with the shared activities of young children who have yet to acquire the skills needed for shared planning agency, it may be important to develop a 'lite' theory of shared intentions, that is, a theory of shared intentional agency aimed at capturing (among other things) what is special about the shared activities of the children. While Bratman may be right that there is some important cognitive, metaphysical and normative continuity between the structures of individual planning agency and those of shared planning agency, it is possible that there is some discontinuity between individual intentional agency and shared intentional agency.

4. The development of joint action

On heavyweight philosophical accounts of shared intentions, having a shared intention demands advanced representational, conceptual and communicational skills and sophisticated forms of reasoning about the complex relationships between each other's actual and inferred goals and intentions. This is problematic given that many developmental findings suggest, on the one hand, that two-year old children already engage in what appears to be intentional joint action and, on the other hand, that children do not yet master at that age the kind of mentalizing skills heavyweight theories require for sharing intentions.

I cannot discuss here the huge literature on the development of mindreading skills in human children. I shall limit myself to three remarks. First, there is consensus that children do not succeed on standard false belief tasks until around 4 years of age (Wellman *et al.* 2001; Wimmer & Perner 1983) and have difficulty before that age understanding intentions and distinguishing them from desires (Astington 1991, 1994; Perner 1991). One should also note that data coming from several recent experimental studies suggest that preverbal human infants as young as 15-month-old (Onishi & Baillargeon 2005), as 13-month-old (Surian, Caldi & Sperber 2007), or even as 7-month-old (Kovacs, Téglás & Endress, 2010) are able to track other agent's false beliefs. While the interpretation of these data is currently the focus of an intense debate,⁶ to my knowledge, no one, however, claims that that they are evidence that infants form explicit representations of other's mental states of the kind needed for flexible theory of mind and practical reasoning. But on heavyweight theories of shared intentions, flexible theory of mind and practical reasoning are needed for coordinating plans and structuring bargaining. Therefore, and this is my second remark, even if the conclusion of the current debate were to be that these data indeed support the attribution to very young children of some mentalizing skills, this wouldn't support the claim that these children have the skills needed to form shared intentions as defined in heavyweight theories.

⁶ See, for instance, Apperly & Butterfill (2009); Baillargeon, Scott, & He (2010), Perner and Ruffman (2005).

Finally, a number of researchers have argued that infants are sensitive to some aspects of goal-directed activity, can attribute goals to actions and can discriminate between intentional and accidental actions (Gergely *et al.* 1995; Csibra 2008; Tomasello and Rakoczy 2003; Woodward 1998; Woodward and Sommerville 2000). However, and this is my third remark, evidence for an early understanding of goal-directed activity isn't yet evidence for an early understanding of intentions *qua* internal mental states. For instance, Csibra and Gergely (2007) propose that the interpretation of observed behaviors of others as goal-directed actions is subserved by teleological reasoning mechanisms, where action outcomes are evaluated with respect to an efficiency principle. In other words, the outcome of an action may, or may not, be seen as its goal, depending on an assessment of the relative efficiency of the action performed to achieve it within the situational constraints given. While there is evidence that infants do engage in teleological reasoning when they observe actions and thus in a sense understand goals, it is also clear that the goals that are thus understood are not mental items but rather states of the world, namely in terms of relations between actions and outcomes that meet certain efficiency constraints in the situation in which this action is produced.

The fact that young children's mentalizing abilities are quite limited would be of little consequence to the issue we are interested in if these same children did not engage in intentional, cooperative joint action. What is the evidence that they do?

In her review of the developmental literature, Brownell (2011) describes several stages in the early development of joint action. During the first year of life, infants already participate in joint action with adult partners who structure the interaction and scaffold the infant's participation. This occurs first in dyadic play and face-to-face interaction. By the end of the first year, infants can initiate and manage this form of joint action with adults and contribute actively to the creation and maintenance of coordinated activity. As pointed out by Brownell, the goal of these social games between young children and their adult partners appears to be affiliation rather than some external, instrumental goal. The end of the first year also sees the emergence of triadic interaction, a form of joint action in which partners act together on objects and events external to the dyad, which is initially structured and led by the adult partners. During the second year of life, children become progressively more adept and active partners in joint activities. By their second birthdays they display skillful coordination even in novel situations,

without the support of familiar goals or routines, and with peers as well as with adults. The studies that first identified this general pattern of development were based on the longitudinal studies that followed infants from 6 to 18 months of age (Bakeman & Adamson, 1984; 1988) or infants between 12 and 24 (Hay, 1979) and observed them playing with their mothers and a set of toys. While these studies could measure the progression of coordinated play around toys and other objects, these informal settings and the routinized and highly practiced character of the activity made it difficult to assess the extent to which infants actively contributed to the coordination.

Warneken, Chen and Tomasello (2006) devised an experimental protocol designed to more accurately assess 18- and 24-month old children's skills in coordinating their actions with those of an adult partner in cooperative activities, including their attempts to regulate the partner's actions during interruptions. They developed four tasks: two problem-solving tasks and two social games with either complementary or parallel roles. For instance, the social game with complementary roles was played by one person sending a wooden block down one of two tubes from the upper side and the other person catching it at the other end with a tin can that made a rattling sound. The two experimenters first demonstrated the task once, and after the demonstration the child was given the chance to perform the task with one experimenter as the partner. During the third and fourth trial, the experimenter acting as partner interrupted the game during 15 seconds before resuming it. The experimenters rated the level of temporal and spatial coordination of the child with their partner and analyzed their behavior during the interruption period. This study yielded three main findings. First, children at 18 and 24 months of age were able to cooperate with an unfamiliar adult partner in a variety of unfamiliar tasks that require the joint activity of two people. Second, the ability to coordinate with the partner significantly improved between children at 18 and 24 months of age. While 18-month old children behavior was rated as predominantly uncoordinated, the older children operated at medium levels of coordination in tasks with parallel as well as with and complementary roles. The third finding concerned the children's behavior during periods in which the partner interrupted his activity: children of both age groups actively communicated to the adult in an obvious attempt to request his reengagement although, while the frequency of communicative acts was similar in both age groups, verbalizations tended to be more frequent among the 24-

month-olds. These findings corroborate earlier results that found a major shift in coordinated activities just before age 2 and extend them in showing that these coordination skills are not restricted to familiar activities performed with familiar partners. They also suggest that a motivation to engage in cooperative activities, as evidenced by attempts to reengage a partner who interrupted the interaction may already be present in children before they master the skills necessary for active coordination.

Thus, as Brownell explains, it seems reasonable to think that infants initially learn about cooperation by participating in joint actions structured and scaffolded by competent adult partners and gradually take more control until they can initiate and manage their contributions themselves and around age two become autonomous collaborative contributors to joint action. There is evidence of this also in interactions with peers, where coordinated actions are infrequent in the second year of life, but present a sharp increase between 24 and 28 months (Holmberg 1980; Eckerman *et al.* 1989)

There is also evidence that, between two and three years of age, young children come to understand and act with respect to the normative dimensions of collaborative activities, including an understanding of explicit agreements, conventional game rules, and commitments to others in joint activities (Gräfenhain *et al.* 2009; Hamann *et al.* 2012; Rakoczy *et al.* 2008; Warneken *et al.* 2006; Warneken & Tomasello 2007). For example, three-year-olds remain committed to following through in a joint activity until both partners have received their rewards, even when one partner, by accident, gets hers first (Hamann *et al.*, 2012). In addition, when three-year-olds form explicit joint commitments with a partner verbally, they expect the other to live up to them in a way that they do not without such a commitment and they themselves feel the need to acknowledge or apologize if they must break off from an explicitly formulated joint commitment with a partner (Gräfenhain *et al.*, 2009). This suggests that three year-olds view their own and the partner's action as being part of a collaborative activity with a joint goal, and thus that their joint actions meet the joint goal requirement.

What is less clear is whether two-year old also view their own and the partner's action as being part of a collaborative activity with a joint goal rather than viewing their partners as mere social tools instrumental in bringing about their own individual goal.

There is of course evidence that very young children are not purely selfish creatures and exhibit pro-social behavior early in life. Human infants as young as 14 to 18 months of age help others attain their goals – for example, by helping them to fetch out-of-reach objects or opening cabinets for them – and do this irrespective of any reward from adults (Warneken & Tomasello 2007). They also engage in other types of pro-social behavior include comforting others, sharing valuable resources and providing useful information for others (Warneken & Tomasello 2009). Furthermore, even very young children are not completely indiscriminate cooperators. For example, infants show preferences for individuals on the basis of common language (Kinzler, Dupoux, & Spelke 2007) and past helpful behavior (Hamlin, Wynn, & Bloom, 2007; Dunfield & Kuhlmeier 2010), and they avoid helping individuals with harmful intentions (Vais, Carpenter & Tomasello 2010).

However, the fact that very young children engage in pro-social behavior is no guarantee in itself that when acting jointly, they view their own and the partner's action as being part of a collaborative activity with joint goals or simply view their partners as social tools whose participation is needed to achieve their own individual goal. For instance, it may be argued that in the study by Warneken, Chen and Tomasello (2006) described above children attempted to reengage their recalcitrant partner because they needed his participation to continue their own individual action rather than because they considered him as a collaborative partner in joint activity. In a more recent study with 27-month old children, Warneken, Gräfenhain, and Tomasello (2012) addressed this issue using 21- and 27-month-old children's responses to the interruption of the activity by the partner in games in which they either did or did not physically need the partner to reach a concrete goal and where the partner stopped interacting, either because she was unwilling to continue or unable to continue. Their reasoning was that on the social tool hypothesis, irrespective of the experimenter being unwilling or unable to continue, attempts to support the partner should be more likely in tasks with causally related actions because only in these tasks is the other person's participation necessary for an individual goal fulfillment. In contrast, on the collaborative partner hypothesis, the reverse pattern should obtain: children should be sensitive to the partner's intention and thus respond differently to interruptions depending on whether the partner does so because she is unwilling or unable to continue, but reengage their partner to the same

extent whether or not her participation is required for individual goal attainment. They found that, congruent with the collaborative partner hypothesis, children of both age groups encouraged the recalcitrant partner equally often whether she was or was not physically needed for goal attainment, but did so more often when the partner was unable to continue than when she was unwilling to continue.

Altogether, these developmental studies provide evidence that young children are not only able to engage in joint activities with others, but view these interactions as genuinely collaborative activities with joint goals. The study by Warneken, Gräfenhain, and Tomasello (2012) suggests that this ability is already present in two-year old children, even before they develop an understanding of the explicitly normative aspects of collaborative activities.

In a nutshell, two-year old children appear to engage in joint actions that meet the joint goal requirement, and yet to lack the kind of cognitive sophistication that would be needed for them to share intentions in the way envisaged by heavyweight theories of shared intentions. What is needed therefore is a less demanding theory of shared intentions. Theories of team agency steer us, I think, in the right direction. To them I now turn.

5. Team Agency

Theories of team agency were developed by economists Sugden (1993, 2003) and Bacharach (2006) in response to problems that arise in classical game theory. Orthodox versions of game theory assume that the players of a game are ideally rational agents and have perfect information: they maximize expected utility, given the expected behavior of others, and they have common knowledge of the game itself and of the rationality of other players. One central motivation for theories of team reasoning is that there are games, such as the Prisoner's Dilemma or the Hi-Lo game, that create problems for classical game theory⁷. In these games, there exists some strategy that is

⁷ In the Prisoner's Dilemma, players have a choice between either defecting or cooperating. The payoff matrix is constructed in such a way that *defect* is the dominant strategy each player (i.e., regardless of what the opponent does, *defect* earns a higher pay-off than *cooperate*). On orthodox game theory, it will therefore be rational for each player to choose defect. Yet, the players would both be better off, if they had both chosen *cooperate*. A substantial number of people see that, since in experiments in which people play the Prisoner's Dilemma for money, anonymously and without repetition, between 40 and 50 % of the participants choose *cooperate* (Sally 1995). In the game of Hi-Lo, players must choose between two

arguably rational and that many people adopt in real life, but which can't be explained or predicted as rational by classical game theory. What team agency theorists think is wrong with the orthodox game theory is its focus on individual rationality. The key move they propose consists in replacing the question "What should I do?" asked separately by each individual, with the question "What should we do?" If instead of reasoning as separate individuals, the players reason as members of a team, then certain strategies that orthodox game theory either considers as irrational or cannot justify as rational will be shown to be collectively rational. To articulate this insight, a theory of team agency should do two things: (1) explain how people come to think of themselves as members of a team and ask the question "What should we do?" rather than the question "What should I do?" and (2) explain how people proceed to answer the question "What should we do?" and characterize the patterns of inference they use (or should use) when reasoning as a team. Roughly, as Bacharach puts it, "somebody 'team reasons' if she *works out the best possible feasible combination of actions for all the members of her team, then does her part in it*" (2006: 121). What a theory of team reasoning does is refine and spell this out in a number of ways, characterizing the inference schemas that capture the modes of reasoning involved in team reasoning.⁸ While team agency theorists tend to agree on how team reasoning should be characterized, they differ in their hypotheses about how groups or teams are formed. My primary interest here will be with Bacharach's view that team formation and therefore team agency is the result of framing.

A frame is a set of descriptors used when thinking about a situation. To take a trivial example, thinking of a glass as half-empty or as half-full are alternative ways in which one can frame a given situation, and whether one uses one frame or the other may have important consequences for how one behaves with respect to this situation. For Bacharach, whether an agent identifies herself as a member of a group or not is a matter

actions, *a* and *b*. They receive something only if they both choose the same, but they get more if they both choose *a* than if they both choose *b*. This puzzle is, if anything, even more puzzling than the first one. It is intuitively obvious that the rational choice for both players is *a*. Yet, orthodox game theory has no explanation of what makes the choice of *a* rational. All the theory says is that if either expects the other to play *a*, then *a* is the rational thing to do, but also that if either expect the other to play *b*, then *b* is the rational thing to do.

⁸ See, for instance, Gold & Sugden (2007, 2008) for a presentation and discussion of those reasoning schemas.

of what frame she uses to represent herself and the agents with whom she is interacting. The frame she uses will in turn determine the way in which she will reason about what to do. Bacharach takes group identification to be a basic human propensity. His concern is with identifying conditions for the production of group identification.

To do that, he relies on theories of group identification developed in social psychology. According to Brewer and Garner (1996), people don't have a single self-concept, but rather a range of self-representations falling within three main categories: personal, relational and collective. The personal self is a self-conception as having a unique, differentiated identity. The relational self is a self-conception derived from connections and role relationships with significant others. The collective self is a self-conception defined in terms of membership in social categories or groups. Self-conception is self-framing, and like other forms of framing has characteristic instability and context-dependence. Psychologists have identified a number of conditions that tend to produce group identification, including belonging to the same social category (e.g., being a woman, a philosopher, a Parisian), to the same *ad hoc* category (e.g., being born on the 1st of June), face-to-face contact, "we" language, shared experience (e.g., being an air crash survivor), common interests, common fate, interdependence, and the presence of a competing outside group (e.g., analytic vs. continental philosophers).⁹ Whether a situation promotes group identification and with which group depends on whether it presents some of these properties – let me call them 'group cues' – and whether they are salient enough to prime the corresponding group frame.

For Bacharach, the use of a group frame engages both an agency transformation and a payoff transformation. In identifying as a member of a group, an agent conceives this group as a unit of agency acting in pursuit of some group-goal. The payoff transformation that accompanies agency transformation involves ranking action profiles according to some group-utility function rather than in terms of the individual utility of the agent's own component of these action-profiles.¹⁰

⁹ See Bacharach (2006, chapter 2) for a review of social psychology work on group identification and Hindriks (in press) for an assessment of the empirical adequacy of the team-agency theorists' appeal to group identification in explaining cooperation.

¹⁰ This aspect of team-agency theories remains under-developed. Bacharach (2006) proposes that action profiles should be ranked according to a Paretian criterion, while Sugden (2003) suggests that they should

Let me make three remarks on Bacharach's framing approach. First, on his view, the agency transformation induced by a group frame does not carry any ontological commitments to supra-individual entities; rather it is, so to speak, an intra-individual transformation of agency, i.e., a matter of individuals conceiving of themselves as members of a group. Similarly, team reasoning is reasoning conducted by individual agents that take themselves to be members of groups. Second, Bacharach insists that the use of a frame is not a matter of choice. Rather group identification as a framing phenomenon is automatically induced by involuntary psychological mechanisms, to which the concept of rationality does not apply.¹¹ Engaging in team reasoning rather than in individual reasoning is therefore not a matter of rational choice. Third, on Bacharach's framing approach, common knowledge of group identification and team reasoning is not a prerequisite for group identification and team reasoning. This is a consequence of the fact that, on Bacharach's view, neither is a matter of (rational) choice.¹² Rather, within this framing approach, the order of explanation is reversed. We do not group identify and team reason because we expect or know that others also group identify and team reason; rather it is because we team reason and group identify that we form expectations about how other members of the group will reason and what they will do.

We will have to reconsider Bacharach's claim that the adoption of a frame is never a matter of choice. But first let us see how we can avail ourselves of Bacharach's team-agency framework to propose a lite characterization of shared intentions.

6. *Shared intentions lite: their scope and limits*

Bacharach (2006: 137-141) gave some indications that in his view shared intentions could not be, *pace* Bratman, constructed out of individual materials but could be arrived

be ranked by summing up the payoffs of the team members. See Hakli et al. (2010) for a discussion of group utility functions and group preferences.

¹¹ In contrast, according to Sugden's assurance theory (Sugden 1993), the purpose of the theory of team reasoning is to explain how people cooperate for mutual advantage. It follows that Sugden's agents will endorse team reasoning only if they have assurance that other team members also endorse team reasoning. Assurance therefore functions as a rational precondition for team reasoning on Sugden's account.

¹² This, of course, is not to say that adopting one frame over another in given situations cannot be externally or evolutionary rational (be to the agent's benefit and enhance fitness). Indeed, in his book Bacharach surveys a body of evolutionary theory that provides evidence that humans have evolved to be cooperators and that the psychological mechanisms that support group identification are the proximate mechanisms that make cooperation possible.

at by the route of team-reasoning in a way that would explain the sense of collectivity that accompanies them. Unfortunately, he died before he had a chance to fully articulate these insights. In this section, I try to exploit, and extrapolate from, his claims in order to define a 'lite' notion of shared intention.

For Bacharach, group agency is a matter of individuals conceiving of themselves as members of a group and engaging in team reasoning. Yet, for there to be an actual shared intention, several agents must engage in team reasoning. Thus, a shared intention would have to be characterized as follows (to keep the formulation simple, I consider a two agent case):

Two persons P1 and P2 share an intention to A, if:

(i) each has a self-conception as a member of the team T, consisting of P1 and P2 (*collective self-framing*);

(i') each believes (i) (*group identification expectation*);

(ii) each reasons that A is the best choice of action for the team (*team reasoning from a group viewpoint*); and

(iii) each therefore intends to do his part of A (*team reasoning from an individual viewpoint*).

I start with some general remarks on this approach and then examine each of these conditions in turn.

First, as Bacharach himself point out, shared intentions arrived at through group identification and team reasoning have features that would explain 'the sense of collectivity' that Searle (1990) thinks accompanies shared intentions:

The nature of team reasoning has this effect, because the team reasoner asks herself not 'What should I do?' but 'What should we do?' So, to team reason, you must already be in a frame in which first person plural concepts are activated. The priming conditions for team reasoning have these effects, because [...] team reasoning, for a shared objective, is likely to arise spontaneously in an individual who is in the psychological state of group-identifying with a set of interdependent

actors; and to self-identify as a member of a group essentially involves a sense of collectivity. (Bacharach, 2006: 141).

Another important commonality between Searle and Bacharach is that for both shared agency remains vested in individuals. As already remarked, agency transformation induced by a group frame does not carry any ontological commitments to supra-individual entities. In contrast to Searle's, however, this Bacharach-inspired approach dispenses with we-intentions as *sui generis* mental states. Instead, it appeals to notions of group identification and team reasoning, but in so doing it can't be accused of conceptual or metaphysical extravaganza, since group identification itself is a well-established phenomenon and team reasoning is simply a matter of using certain patterns of inference.

Here, the first step on the route to shared intentions involves group identification and group identification essentially involves a sense of collectivity. As a result of this feature, shared intentions arrived at by this route will be guaranteed to meet the joint goal requirement. This is in contrast to the complex set of conditions, and the accompanying set of sophisticated cognitive skills, that are needed on Bratman's account of shared intentions to meet this requirement.

One important advantage of this framing route to shared intentions is that it allows shared intentions to emerge in situations where the preconditions for the formation of heavyweight shared intentions are not met. First, it allows shared intentions to be formed in situations where agents cannot communicate or influence each other in other ways, as is the case for instance in one-shot versions of the Prisoner's Dilemma and the Hi-Lo game. This is in contrast to Bratman's theory of shared intention. As team reasoning theorists have pointed out (Bacharach 2006: 138 sq; Bardsley 2007; Gold and Sugden 2007, 2008), in Bratman's account, the decisions and actions of the agents are governed by classical (i.e. individualist) canons of rationality. Unless they can influence each other's intentions through their own intentions or actions, they won't be able to rationally generate determinate expectations about others' actions. Thus, in situations where agents are unable to influence each other, there is no way they can rationally decide to cooperate.

Second, and relatedly, it allows shared intentions to be formed by agents who lack sophisticated mentalizing skills. Group identification involves sensitivity to certain features of situations, including the presence of other intentional agents that are potential competitors or cooperation partners. It therefore presupposes a capacity to detect intentional agency and to represent others as intentional agents. However, the contents of the participatory intentions members of team form when they team reason need not make reference to either their own intentions and the intentions of the other participants. In addition, this account doesn't require common knowledge of group identification but only expectation thereof. Agents who engage in team reasoning and expects others to do the same may well be led to form further expectations as to what others will do, but forming expectations about others' actions is not the same thing as ascribing intentions to them and does not require the kind of cognitive and conceptual skills intention ascription requires.

This not to say that a team-agency route to shared intentions does make other theories of shared agency superfluous or redundant. Just as the lite version of a software application typically consumes less computer resources than the full version but may have limited functionality, a lite theory of shared intention may be less cognitively and conceptually demanding but have limitations. I close this paper with a brief discussion of some of these limitations. To see what they might be, we need to consider what it may take to satisfy each of the conditions in the lite account of shared intentions.

On the team-agency approach, condition (i) is realized through group framing: agents think of themselves as members of a group or team as a result of a psychological framing process. Whether a group frame is induced in a given situation depends on what features this situation presents and whether these features are sufficiently salient to be perceived and make the availability of the frame high. Thus, it is likely that two agents, endowed with a similar psychological makeup and placed in a situation where 'group cues' are highly salient, will both adopt the same group frame. This is the reason why group identification expectation was introduced as condition (i'). It isn't independent from condition (i), but rather constitutes a default assumption given the nature of the psychological processes that are supposed to automatically induce group-identification. In other words, had 'non-group' cues be present and salient enough, the group-frame would not have been induced in the first place. But what about situations where group

cues are not salient or where the cues are ambiguous? While Bacharach's idea that group identification is a matter of framing, hence that team reasoning can be spontaneous, is an important insight, I am not sure we should follow him in thinking that the adoption of a frame can never be a matter of choice. It seems in principle possible to recognize the ambiguous character of a situation and hence the possibility that it be framed in different ways. Once we do, several questions may arise: Is one way of framing the situation better than the other? How are other agents likely to frame it? What can I do to make them more likely to adopt one frame? Once these questions are on the table, the framing route is not guaranteed to lead to the formation of shared intentions and agents may need to use other routes, such as Bratman's or Gilbert's, to convey to potential collaborative partners their willingness to engage in joint action with them.¹³

Condition (ii) pertains to the choice of a group goal and involves team reasoning from the group viewpoint. The basic inferential principles used in team reasoning are analogous to the inferential principles used in individual practical reasoning: we must choose between different courses of actions with different outcomes and we should choose the course of action whose outcome we prefer. The main difference here is simply that outcomes are evaluated in terms of group preferences rather than in terms of individual preferences. However, this story is idealized in at least two ways. First, it is implicitly assumed that the alternative action courses among which the agents have to choose are already defined. In real life, however, we often know what we would like to achieve without having a clear idea what we should do to achieve it or, when faced with a novel situation, have yet to figure out what the possible alternative courses of action are. A different variety of reasoning skills may have to be employed to reach the stage where the options are clear. Second, as pointed out by Hakli, Miller and Tuomela (2010), Bacharach's theory takes group preferences for granted or at least it takes for granted that agents are working with the same group preferences. Yet, it may not be obvious

¹³ Note that Bacharach was in a sense well aware of these problems since he developed a theory of circumspect team reasoning in order to deal with such cases. Bacharach (1999) introduces the idea that, in coming to frame a situation as a problem for us, an individual also gains some sense of how likely it is that another individual would frame it in the same way. In circumspect team reasoning, an individual who group-identifies will then maximize the expected value of the group pay-off function given the probabilities that other group members identify or fail to identify to the group. There seems to be some tension, however, between Bacharach's theory of circumspect team reasoning and his claim that group-identification and team reasoning are a matter of framing and hence cannot be a matter of choice. It seems that if one can be aware that a situation can be framed in different ways, which is why circumspection is possible and needed, then it should also be possible to decide which frame to adopt.

what utility values individuals associate to various outcomes or how individual preference orderings relate to collective preference orderings. It may be that in some cases the same situation that triggers a group frame also makes salient a particular group preference function or, in other cases, it may be that, although different agents operate with different preference functions, these different functions rank the same outcome highest. Yet, they will surely also be cases where these group preferences functions are in conflict and where bargaining is called for.¹⁴

Finally, condition (iii) in the team-agency approach pertains to the formation by a member of a team of an intention to do his or her part in the team action. This intention is formed by a process of team reasoning from an individual viewpoint and moves the agents from the premises that a certain team action is best for the team and that a certain component of it is her part in it, to the intention to contribute this part to the joint action. Once again, there are some complexities hidden behind this simple formulation of what is involved in the formation of an agent's intention to contribute her part. In everyday life, a team action is often chosen before the exact division of labor is worked out among agents. Housemates may decide to throw a party together, and know that this will involve sending invitations, clearing aside the furniture in the living room, and providing music, drink and food, without the repartition of roles in accomplishing these tasks being fixed in advance. Yet, it won't do if everyone buys food but no one takes care of the other things that need to be done. The housemates need to coordinate their plans and actions. In some cases, the action affordances in the situation will suffice to ensure coordination. If our joint goal is to move aside a heavy table and you happen to stand near one end of the table and I near the other, our respective spatial positions will dictate what our parts are. In other cases, the individual capacities of the agents will also impose some constraints on role assignments. Josh, who is tone deaf, will not be put in charge of the party music, and Sacha, the teetotaler, will not be asked to buy the drinks. In other cases yet, the role assignments will be obvious: Boris will sing the tenor part of the duet and Irina the soprano part. Yet, there will also be more complex cases where negotiation and bargaining of the kind envisaged by Bratman will be needed before roles are assigned and agents can form relevant intentions concerning their parts in the joint action.

¹⁴ For a discussion of group preferences, see Tuomela (2007, chapter 7).

7. Conclusion

In this paper, I tried to show how using the framework of Bacharach's framing theory of team agency, one could develop a lite account of shared intentions. The main virtue of this account is, in my view, that it allows us to understand how agents acting jointly can view their own and the partner's action as being part of a collaborative activity with joint goals, without needing to deploy sophisticated conceptual and mindreading skills. This account of shared intentions is thus applicable to the joint actions performed by young children who lack these skills. It is also applicable, in many cases, to joint actions performed by human adults, for even though they have those skills, they do not always marshal them when they engage in intentional joint action.

Yet, this account is not a panacea and has limitations of its own. First, we can also form shared intentions and perform joint actions in situations where cues to group framing are weak or not readily available. Second, we do not always get group preferences for free; they may have to be constructed. Third, it is not always out in the open what the action alternatives are and how parts should be assigned. This account should therefore be supplemented with accounts of how action alternatives and group preferences can be worked out by agents and how roles are distributed. Finally, an account of shared intention should tell us not just what it takes to engage in joint action, but also what it takes to ensure it is successfully performed. Team agency theory allows us to understand how it is possible for people to form joint goals and intentions to do their parts in achieving these goals. Yet, it has very little to say about how in many real-life situations people succeed in coordinating their actions towards a joint goal. To answer this question, we must appeal to resources that go beyond what the theory has to offer.

Acknowledgments I would like to acknowledge valuable comments and suggestions on earlier versions of this paper from Stephen Butterfill, from participants at the First PML Conference in Stockholm in September 2011, and from two anonymous referees.

References

Apperly, I.A., and S.A. Butterfill. 2009. Do humans have two systems to track beliefs and belief-like states? *Psychological Review* 116(4): 953–970.

- Astington, J. (1991). Intention in the child's theory of mind. In D. Frye & C. Moore (Eds.), *Children's Theories of Mind: Mental States and Social Understanding*, Hillsdale: Lawrence Erlbaum Associates.
- Astington, J. (1994). *The Child's Discovery of the Mind*. Cambridge, Mass: Harvard University Press.
- Bacharach, M. (1999). Interactive team reasoning: a contribution to the theory of cooperation. *Research in Economics*, 53, 117-147.
- Bacharach, M. (2006) *Beyond Individual Choice: Teams and Frames in Game Theory* (edited by N. Gold and R. Sugden). Princeton, Princeton University Press.
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14, 110-118.
- Bakeman, R., and L.B. Adamson. 1984. Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, 55(4): 1278-1289.
- Bakeman, R., and L.B. Adamson. 1986. Infants' conventionalized acts: Gestures and words with mothers and peers. *Infant Behavior & Development* 9: 215-230.
- Bardsley, N. (2007) On collective intentions: collective action in economics and philosophy. *Synthese*, 157, 141-159.
- Blomberg, O. (2011). Socially extended Intentions-in-action. *Review of Philosophy and Psychology*, 2:335-353.
- Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2), 327-41.
- Bratman, M. E. (1999). *Faces of Intention*. Cambridge: Cambridge University Press.
- Bratman, M. E. (2009). Modest Sociality and the Distinctiveness of Intention. *Philosophical Studies*, 144, 149-165.
- Brewer, M. B., & Kramer, R. M. (1986). Choice behavior in social dilemmas: Effects of social identity, group size, and decision framing. *Journal of Personality and Social Psychology*, 50: 543-549.
- Brownell, C.A. (2011). Early Developments in Joint Action. *Review of Philosophy and Psychology*, 2: 193-211.
- Butterfill, S. (2011). Joint action without shared intention. Manuscript.
- Butterfill, S. (2012). Joint Action and Development. *The Philosophical Quarterly*, 62, 246: 23-47.
- Butterfill, S. & Sebanz, N. (2011). Joint action: What is Shared? *Review of Philosophy and Psychology*, 2: 137-146.
- Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition*, 107(2), 705-717.
- Csibra, G. & Gergely, G. (2007). 'Obsessed with goals': Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica*, 124: 60-78
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Oxford University Press.
- Dunfield, K. A., & Kuhlmeier, V. A. (2010). Intention-Mediated Selective Helping in Infancy. *Psychological Science* 21, 4: 523-527.
- Eckerman, C. O., Davis, C. C., & Didow, S. M. (1989). Toddlers' emerging ways of achieving social coordinations with a peer. *Child Development*, 60, 440-453.
- Gergely, G., Nadasky, Z. Csibra, G & Biro, S. (1995). Taking the Intentional Stance at 12 Months of Age. *Cognition*, 56, 165-193.
- Gilbert, M. (1992). *On social facts*. Princeton: Princeton University Press.
- Gilbert, M. (2009). Shared intention and personal intentions. *Philosophical Studies*, 144, 167-187.
- Gold, N. and R. Sugden (2007) Collective Intentions and Team Agency. *Journal of Philosophy*, 104(3), 109-37.
- Gold, N. and R. Sugden (2008) Theories of Team Agency. In F. Peter and S. Schmidt (Eds.) *Rationality and Commitment*. Oxford: Oxford University Press.
- Gräfenhain, M., Behne, T., Carpenter, M., & Tomasello, M. (2009). Young Children's Understanding of Joint Commitments. *Developmental Psychology*, 45(5), 1430-1443.
- Hakli, R., Miller, K, & Tuomela, R. (2010). Two kinds of we-reasoning. *Economics and Philosophy*, 26, 291-320.
- Hamann, K., Warneken, F. & Tomasello, M. (2012), Children's developing commitments to joint goals. *Child Development*, 83(1), 137-145.
- Hamlin, K.J., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, 450, 557-560.
- Hay, D.F. 1979. Cooperative interactions and sharing between very young children and their parents. *Developmental Psychology*, 15(6): 647-653.
- Helm, B. W. (2008). Plural agents. *Nous* 42(1): 17-49.
- Hindriks, F. (in press). Team Reasoning and Group Identification. *Rationality and Society*.
- Holmberg, M. 1980. The development of social interchange patterns from 12 to 42 months. *Child Development* 51: 448-456.

- Jeannerod, M. (1997). *The cognitive neuroscience of action*. Oxford: Blackwell.
- Kinzler, K.D., Dupoux, E., & Spelke, E. (2007). The native language of cognition. *Proceedings of the National Academy of Sciences, USA*, 104, 12577–12580.
- Knoblich, G. and Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions of the Royal Society B*, 363: 2021–2031.
- Knoblich, G., Butterfill, S., and Sebanz, N. (2010). Psychological research on joint action: Theory and data. In Ross, B., editor, *Psychology of Learning and Motivation*, volume 51. Academic Press.
- Kovács, A.M., Téglás, E., & Endress, A. D. (2010). The Social Sense: Susceptibility to Others' Beliefs in Human Infants and Adults. *Science*, 330: 1830-1834.
- Kutz, C. (2000). Acting Together. *Philosophy and Phenomenal Research*, 61, 1, 1-31.
- Michael, J. (2011). Shared Emotions and Joint Action. *Review of Philosophy and Psychology*, 2: 355–373.
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science*, 308(8), 255-258.
- Pacherie, E. (2011). Framing Joint Action. *Review of Philosophy and Psychology*, 2, 2: 173-192.
- Pacherie, E. & Dokic, J. (2006). From mirror neurons to joint actions. *Journal of Cognitive Systems Research*, 7, 101-112.
- Perner, J. (1991). On representing that: The asymmetry between belief and desire in children's theory of mind. In D. Frye & C. Moore (Eds.), *Children's Theories of Mind: Mental States and Social Understanding*. Hillsdale: Lawrence Erlbaum Associates.
- Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, 308, 214–216.
- Rakoczy, H., Warneken, F., & Tomasello, M. (2008). The sources of normativity: young children's awareness of the normative structure of games. *Developmental Psychology*, 44(3), 875-881.
- Sally, D. (1995). Conversation and Cooperation in Social Dilemmas: A Meta-Analysis of Experiments from 1958 to 1992. *Rationality and Society*, 7, 58–92.
- Searle, J. (1990) Collective Intentions and Actions. In P.Cohen, J. Morgan, and M.E. Pollack (Eds.), *Intentions in Communication* (pp. 401-416). Cambridge, MA: Bradford Books, MIT Press.
- Seemann, A. (2011) Joint Motor Action and Cross-Creature Embodiment. *Review of Philosophy and Psychology*, 2: 279–301.
- Sugden, R. (1993) Thinking as a Team: Towards an Explanation of Nonselfish Behavior. *Social Philosophy and Policy*, 10: 69-89.
- Sugden, R. (2000) Team Preferences. *Economics and Philosophy*, 16: 175-204.
- Sugden, R. (2003) The Logic of Team Reasoning. *Philosophical Explorations*, 6: 165-81.
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of Beliefs by 13-Month-Old Infants. *Psychological Science*, 18(7), 580-586.
- Tollefsen, Deborah (2005). Let's pretend: Children and joint action. *Philosophy of the Social Sciences*, 35(75): 74–97.
- Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language*, 18(2), 121-147.
- Tuomela, R. (2007). *The Philosophy of Sociality: The Shared Point of View*. New York: Oxford University Press.
- Vaish, A, Carpenter, M, & Tomasello, M. (2010). Young Children Selectively Avoid Helping People With Harmful Intentions. *Child Development*, 81, 6: 1661–1669
- Warneken, F. & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, 11, 271-94.
- Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative Activities in Young Children and Chimpanzees. *Child Development*, 77, 3, 640 – 663.
- Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science*, 15(1), 54-61.
- Wellman, H. Cross, D., & Watson, J. (2001). Meta-Analysis of Theory of Mind Development: The Truth About False-Belief. *Child Development*, 72: 655–84;
- Wimmer, H., & Perner, J. (1983). Beliefs About Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. *Cognition*, 13: 103–28;
- Woodward, A. L. (1998). Infants Selectively Encode the Goal Object of an Actor's Reach. *Cognition*, 69, 1-34.
- Woodward, A. L., and Sommerville, J. A. (2000). Twelve-Month-Old Infants Interpret Action in Context. *Psychological Science*, 11(1), 73-77.